KATHOLIEKE
UNIVERSITEIT
LEUVEN

DEPARTEMENT TOEGEPASTE ECONOMISCHE WETENSCHAPPEN

RESEARCH REPORT 0253

PERFORMANCE OF BELGIAN MUTUAL FUNDS: DO SIZE AND MOMENTUM MATTER?

by
J. CROMBEZ
L. DE MOOR
P. SERCU
R. VANPEE

D/2002/2376/53

Performance of Belgian Mutual Funds: do Size and Momentum matter?

John Crombez*, Lieven De Moor†, Piet Sercu‡ and Rosanne Vanpee§

First draft: November 2002
Preliminary — comments very welcome

*Universiteit Gent, Department Financial Economics, john.crombez@rug.ac.be

†KU Leuven, Departement TEW, lieven.demoor@econ.kuleuven.ac.be

‡KU Leuven, Graduate School of Business Studies, piet.sercu@econ.kuleuven.ac.be. Corresponding author; Naamsestraat 69, B-3000 Leuven

§WH Selfinvest, Luxemburg, rosannevanpee@hotmail.com

Abstract

In light of the well-known empirical failures of the one-factor CAPM, mutual-fund performance evaluation should venture beyond the one-factor type of performance analysis. In this paper we introduce momentum and size factors into the picture, and evaluate the performance of a large set of equity funds managed in Belgium. There is a fairly strong exposure to the small-firm effect, but the evidence of momentum chasing is less clear-cut and, if anything, seems to be negative. As in other studies, the average fund underperforms. Nor do we find any instances of excess performance when grouping funds by management company.

Keywords: Mutual Fund Performance, Size, Momentum.
JEL-codes: G32, G34.

Performance of Belgian Mutual Funds: do Size and Momentum matter?

Introduction

To date, the most popular measures of mutual-fund performance still are the Sharpe ratio (Sharpe, 1964) and Jensen's alpha (Jensen 1968, 1969)—despite the empirical criticisms of the one-factor asset-pricing model that underlies these measures. In academe, and increasingly also in industry, the observed failings of the one-factor model have led to more sophisticated evaluation procedures that, however, still retain the spirit of the original: multi-factor Capital Asset Pricing Models, Arbitrage Asset Pricing Theory, or Stochastic Discount Factor models, all reviewed in *e.g.* Sercu (2000). These models can also be used to identify the style of the fund—that is, the niches the fund shuns or over-invests in, for example large stocks, or distressed firms, and so on—provided that suitable style-indices are introduced as yardsticks.

This paper illustrates how the addition of even just two factors considerably changes the picture, fund by fund. The exposures to the factors imply that, in a larger sample (or one where factor mean returns do not happen to be close to zero) also the average performance across fund will be substantially affected. In what follows we first review, very briefly, the competing one- and multifactor models and the resulting yardsticks. We then present the results from our three-factor analysis and contrast them with the one-factor counterparts (Section 2). Section 3 concludes.

1 Yardsticks for Mutual-Fund Performance

The notation we adopt is the following. R stands for the return, that is, the simple percentage change in the value of the asset or portfolio over a time period. Subscript f to R refers to a particular mutual fund. Subscript 0 refers to the zero-risk asset (a T-bill, or a high-quality CD or bank deposit). A tilde (˜) over the R (or over another symbol) means that the variable is random (or "risky"). $\mathrm{E}(\tilde{R}_f)$ is the unconditional expected (risky) return on the fund over an unspecified period—the unconditional probability-weighted average of all possible outcomes. Expected returns can and do change over time. Thus, $\mathrm{E}_{t-1}(\tilde{R}_{f,t})$, refers to the return over a period t (say, March 2003), as it is (or was) expected at time $t-1$, the beginning of that period, on the basis of all then available information. Thus,$\mathrm{E}_{t-1}(\tilde{R}_{f,t})$, is the conditional expected (risky) return on the fund over a specified period—the conditional probability-weighted average of all possible outcomes. We use $\mathrm{var}(\tilde{R}_f)$ to denote the variance, and $\mathrm{std}(\tilde{R}_f)$ (that is, $\sqrt{var(\tilde{R}_f)}$) to denote the standard deviation of the fund's return. Lastly, the excess return of fund f is defined to be $(\tilde{R}_f - R_0)$, that is, its return in excess of the risk-free rate.

We first review the one-factor-CAPM performance measures. Familiarly, the Sharpe ratio is the fund's expected excess return scaled for risk by dividing by its standard deviation, and Jensen's alpha is the mean excess return left unexplained by the asset's market sensitivity or beta:

$$\text{Sharpe Ratio} = \frac{E(\tilde{R}_f) - R_0}{\mathrm{std}(\tilde{R}_f)} \tag{1.1}$$

$$\text{Jensen's alpha} = \mathrm{E}_{t-1}(\tilde{R}_{f,t} - R_{0,t}) - \beta_{f,t-1}\mathrm{E}_{t-1}(\tilde{R}_{M,t} - R_{0,t}) \tag{1.2}$$

where $\beta_{j,t-1}$ is the slope of a regression of $\tilde{R}_{f,t} - R_{0,t}$ on $\tilde{R}_{M,t} - R_{0,t}$. In practice, the alpha is obtained by assuming the beta in (1.2) to be constant over time, running an OLS regression with an intercept α_f added, and testing whether α_f is clearly non-zero (and, preferably, positive).

Both performance measures are based on "mean-variance" portfolio theory and the resulting one-factor Capital Asset Pricing Model (CAPM1). Relative to Sharpe's ratio, Jensen's alpha has the major advantage that it can be logically applied to any individual asset, not just to the investor's entire portfolio. Still, the alpha is far from perfect yet. First, the CAPM1 assumes a mean-variance world, which is problematic both a priori and empirically. The CAPM indeed predicts that differences in expected returns should be explained by differences in beta and by nothing else, but this turns out to be empirically untrue.[1] One major "anomaly" is size: in the past, the smallest stocks have on average outperformed the return predicted on the basis of their beta, and the logarithm of the market capitalization turns out to be a better predictor of mean returns than the OLS beta. Related anomalies are the Price/Earnings and book-to-market factors: the low-P/E stocks outperform what one would have expected on the basis of beta, and so do distressed stocks with a high book-to-market, that is, a high ratio of book value to market value. Another

[1] For a review, see e.g. Fama and French (1992).

anomaly is short-term momentum: short-term winners (that is, stocks that outperformed over the last 3-12 months) tend to go on winning, while short-term losers tend to go on losing. Lastly, and statistically more ambiguous, there may also be a long-term reversal phenomenon: stocks that badly underperformed the market over the past five years tend to provide unusually high returns and vice versa.

Initially these phenomena were viewed by many to reflect market inefficiencies. Yet they could also be rational priced factors. The theoretical existence of such additional factors, beside the market portfolio, was first suggested in the intertemporal CAPM of Merton (1973): expected excess returns should be determined by assets' sensitivities to more portfolios than just the market portfolio if the additional portfolios stand for underlying factors that change the investment opportunity set.[2] Somewhat more pragmatically, however, the existence of the "anomalies" does mean that the portfolio manager may achieve "abnormal" returns (that is, a positive alpha) by simply chasing the anomalies. That is, the manager may simply be investing into small stocks or low-P/E firms, etc. rather than being good at picking individual stocks. Multifactor-based measures of performance do attempt to sort out this effect.

Such multi-factor or multi-index models can also be used to identify the fund's style. In "style analysis", first promoted by Sharpe (1992), one regresses the fund's excess returns (as of now denoted by $\tilde{r}_f \equiv \tilde{R}_f - R_0$) on the excess returns from N market-segment indices. A general style-analysis regression with benchmarks $b_1, ..., b_n$ can be written as

$$\tilde{r}_{f,t} = \alpha_f + \beta_{f,b_1}\tilde{r}_{b_1,t} + \beta_{f,b_2}\tilde{r}_{b_2,t} \cdots + \beta_{f,b_n}\tilde{r}_{b_n,t} + \tilde{\epsilon}_{f,t} \tag{1.3}$$

For instance, the benchmarks used in Sharpe's website are (i) S&P firms with a large market capitalization ("cap") and a high book-to-market; (ii) S&P firms with a large cap and a low book-to-market; (iii) mid-cap firms (that is, the larger non-S&P stocks); (iv) small-cap non-S&P firms; (v) long-term Government bonds (20 years); and (vi) long-term corporate bonds. Regressions like (1.3) tell us, firstly, what kind of stocks the fund has been really after, possibly in contrast to its stated policy. A second purpose of these regressions may be to take into account "significant" differences between returns in various segments of the

[2] In a multi-period setting, investors logically care not just about the mean and variance of wealth at the end of the current period, but also about the conditions under which they will be able to re-invest afterwards (the "investment opportunity set"). Business cycles, or changes in macro uncertainty obviously do change the investment opportunity set; and smaller firms or distressed firms with high book-to-market are more sensitive to business cycles or changes in macro uncertainty.

market.[3] Lastly, one could view the expected-value version of (1.3) as a multi-factor CAPM, and use the intercept α_f as the multifactor generalization of Jensen's abnormal return.

Regarding the choice of factors, US studies usually fall back on the extensive tests of multifactor CAPMs by Fama and French (1996), who conclude that the following three factors do a good job in describing/explaining asset returns: $\tilde{r}_M$, the excess return on the value-weighted market portfolio M; SB, the small-minus-big factor or size factor [return on the smallest stocks, minus return on the biggest stocks]; and HL, the high versus low book-to-market value factor [return on high book-to-market stocks, minus return on low book-to-market stocks]. These factors, however, fail to explain the expected returns on momentum-specialized portfolios. Accordingly, recent studies like Carhart (1997) add a fourth factor that should capture this continuation phenomenon: Mo, the momentum factor, [return on short-term winners, minus return on short-term losers]. For an exploratory study as ours, we dropped the HL factor, which requires hard-to-find accounting data (practical issue). Moreover Crombez (2001) finds that European equity is better priced by models including Mo and SB than models with HL (empirical issue). Thus, our regression is

$$\tilde{r}_f = \alpha_f + \beta_f \tilde{r}_{M,t} + \gamma_f \tilde{SB}_t + \delta_f \tilde{Mo} + \tilde{\epsilon}_{f,t}. \tag{1.4}$$

2 A three-factor evaluation of Belgian Mutual funds.

2.1 Data selection

We use three years of data, January 1999 till December 2001, from Datastream, obtained as follows:

- *Mutual Funds.* We use the (USD) returns of all 322 international equity funds that are sold in Belgium and Luxemburg, are listed as such in the *De Collectieve beleggingen in Belgie 2001, Beleggingsfondsen en Beleggingsvennootschappen*, and have al least twelve months of data. They represent 26 different fund-management firms. Starting from a

[3]Suppose, for instance, that the one-factor CAPM holds but that, during a particular period, the bond market did quite well while stocks did unusually poorly. Regressing bond-fund excess returns on a market excess return that comprises both stocks and bonds would then misleadingly suggest that the bond-fund managers were very clever (their alpha is positive); likewise, stock-fund managers would look undeservedly bad. One could try to handle this by regressing stock-fund returns on a pure stock-market index, and bond-fund returns on a pure bond-market index, but this would be incompatible with the CAPM (which says that covariance with the entire market of stocks and bonds is what matters). It can be shown that including into the regression both segments (the one the fund is going into, and the remaining part of the market) is an appropriate procedure. That, of course, is what style analysis does.

list of funds that existed in the middle of the period, we might miss funds that had gone under in 1999-2000, but in view of the booming market in those years the risk of survival selection bias is small. Anyway, the focus is on the importance of the three-factor evaluation. We report no individual names of funds nor of managing firms: we find little systematic evidence of differential stock-picking abilities across managers, and we want to avoid our numbers being quoted selectively without our warning that any differences are probably unreliable for the purpose of diagnosing the past and, *a fortiori*, predicting the future.

- *Market and Small-Big Returns.* The CAPM logic suggests that we take a worldwide index as the benchmark.[4] Our index is Morgan Stanley Capital International (MSCI), which now covers 95 percent of the aggregate market cap of all stocks. With respect to the size and momentum factors, we take into account the manifest home bias observed by others. Hoping to pick portfolios that best reflect opportunity sets relevant for the European stocks that our funds largely invest in, the SB return factor is constructed from the returns on two European indices, Dow-Jones' STOXX Total Market Index Large Cap and on the DJ STOXX TMI Small Cap. All these indices are total-return indices, that is, they compute values with reinvestment of the dividends. The Large and Small indices are based on free-float caps; DJ STOXX TMI Large consists of firms from the 70 percent biggest-cap segment, and the corresponding Small index from the smallest-but-one 5-percent segment, that is, the firms between the 90th and 95th percentiles of all firms ranked downward by size. The composition of each index is revised quarterly. The returns are taken from Datastream.

- *Momentum.* The momentum factor was constructed as follows. First, we extract the data of 2657 individual stocks from Datastream. Datastream has the drawback that it omits the smaller firms, and also removes all delisted firms from its current files,

[4] Early applications of the CAPM, mostly on US data, consistently use a US index as the basis of the benchmarking exercise. This procedure implicitly views the US as a closed market, with local investors choosing just local stocks and with all local stocks being US-held. While this was not unreasonable for the US in the 70s, for post-1900 European markets it is an untenable assumption: we all hold many foreign stocks, and many domestic stocks are held abroad. In internationally integrated markets, the market should be the world market, as shown in the International CAPM by Sercu (1980). True, for funds dedicated to just one country or one industry, a local or industry index would have generated higher R^2s. Still, the purpose is not high R^2s but a theoretically sound and coherent approach. Equally true, the real exchange rates that should enter as additional InCAPM factors are missing in this study, but evidence that mainstream-currency risk-free assets have different expected returns in the long run is skimpy.

thus potentially creating a severe survival bias.[5] There is nothing we can do against the large-firm bias in Datastream—there is no reason to suspect that a large-firm bias affects momentum anyway—but we do try to tackle the survival bias, as follows. We take firms from Datastream's "market list" until we have covered 80 percent of the market cap. We then look at Datastream's so-called "dead list", compute average market values of those firms, and select vanished firms until we also cover 80 percent of that list. This way we end up with our 2657 stocks. For every month in the sample we then proceed as follows. We compute each stock's return over the past six months, rank all returns, and pick up the 25 percent top and bottom performers. Mo is then computed as the return on an equally weighted portfolio of winners minus the return on an equally weighted portfolio of losers.

2.2 Empirical results (1): stock-picking ability on the whole and by fund manager

Panel A of Table 1 provides a summary of the parameter estimates and their t-statistics. As descriptive non-parametric statistics we show the lowest, midway (or median), and highest numbers. The usual parametric ingredients are also present: the mean, standard deviation, and the t-statistic on the mean, computed as $avg/std * \sqrt{N}$ with $N = 322$ being the number of funds in the current sample. For every parameter we thus have (i) a t-test on the mean of the estimated parameter and (ii) a t-test on the mean of the t-statistics of that estimator. In principle, the t-test on the mean of t-tests is more reliable than the direct t-test on the mean of the parameters, because the former underweighs low-quality estimates and relies relatively more on the more precise numbers.[6] A potential weakness of either t-test is that it assumes independence of the observations in the sample. With respect to selectivity (α_f), one could argue that there could be a downward bias in the standard deviations due to possible commonalities in all funds run by the same managing firm. Thus, both t-tests would be overstating the significance since the standard deviation for the mean would be underestimated. Lastly, we compute the returns for two alternative portfolios of all mutual

[5] If firms are mainly delisted because they go bankrupt or are taken over under distress circumstances, then a data base that contains just survivors would never notice the bad returns on deceased firms and, therefore, overestimate the normal returns.

[6] In practice we find no systematic difference between the two sets of t-tests on averages. Perhaps there is little difference in the precision of the estimates across funds, and/or the tests are rather weak at assessing any quality differences present in the results.

Table 1: 1-factor evaluation: market, SB, and momentum sensitivities and average abnormal returns

$$\tilde{r}_f = \alpha_f + \beta_f \tilde{r}_{M,t} + \gamma_f \tilde{SB}_t + \delta_f \tilde{Mo} + \tilde{\epsilon}_{f,t}.$$

Panel A: Summary statistics for 322 individual evaluations

	α(%)	β	γ	δ	t(α)	t(β)	t(γ)	t(δ)	$\bar{R}^2$
min	-4.089	-0.828	-1.162	-1.020	-3.269	-1.775	-2.647	-5.665	-0.141
med	-0.111	0.996	0.353	-0.050	-0.151	4.797	1.237	-0.340	0.580
max	6.319	4.003	2.157	0.898	4.678	22.704	5.272	3.427	0.955
avg	-0.096	1.080	0.442	-0.059	-0.140	5.082	1.365	-0.411	0.536
std	1.092	0.607	0.486	0.251	0.992	3.501	1.324	1.534	0.241
t(av)	-1.881	31.913	16.334	-4.218	-2.525	26.047	18.505	-4.805	—

Panel B: Statistics for aggregate funds

weight	α(%)	β	γ	δ	t(α)	t(β)	t(γ)	t(δ)	$\bar{R}^2$
equal	-0.261	1.119	0.031	-0.100	-0.598	8.071	0.295	-1.697	0.808
$\text{var}^{-1}_{\alpha_f}$	-0.146	0.929	-0.010	-0.009	-0.451	8.983	-0.122	0.044	0.870

Key to Table 1. *Panel A.* Monthly excess returns, 1999-2001, on 322 Belgian-managed internationally oriented stock funds are regressed on the world market excess return r_m (MSCI), a small-minus-large return factor SB (DJ STOXX Small minus DJ STOXX Large), and a momentum return Mo (the 25% best winners *vs.* the 25% worst losers over the past six months, taken from a sample of over 2500 individual European stocks). β, γ, and δ estimate the fund's sensitivity to these factors. α estimates the average return that is not explained by any of these factors, and is in percent per month. Min stands for lowest estimate, med for Median (the centrally ranked estimate), avg for average, max for highest, std for standard deviation, and $t(av)$ for $(avg/std) \cdot \sqrt{N}$. Note that $std/\sqrt{N}$, on average, underestimates the standard deviation of the mean, so the t-test on averages and on t-tests overstate the significance. *Panel B.* The evaluation is done on a fund of funds, growing from 193 (early 1999) to 322 (end 2001). The first fund-of-funds has equal weighting, in the second the individual funds are weighted inversely proportional to the estimation variance of their alpha in the first-pass regression reported in Panel A. The t-statistics are based on OLS assumptions and ignore the fall in residual risk when the number of investments grows from 193 tot 322; that is, the tests again overstate the significance.

funds, and evaluate each of those fund-of-funds in the same way as genuine funds. The two funds-of-funds differ in the way the individual funds are weighted. In the first portfolio, we weight equally, thus giving each manager an equal importance. In the second we weight each fund inversely proportional to its variance of alpha as obtained from the individual regressions, thus giving more reliable estimates more weight.[7]

Let us turn to the alpha and t-alpha columns in Table 1, the measures of the managers' stock-picking abilities net of costs. In line with about every large-scale study on the mutual-fund industry, we find that the average abnormal return—here the return not explained by

[7]The (impractical but) theoretically superior scheme is to use weights $w = \Omega^{-1}u/(u'\Omega^{-1}u)$ where Ω is the 322 x 322 residual variance-covariance matrix across the regressions and u the unit vector. Weighting by inverse variance, as we do, ignores the off-diagonal elements, which makes the weighting less efficient but does not invalidate the t-statistic.

the general market, the size factor, and momentum—is negative; that is, from the point of view of the mutual-fund shareholder, the typical fund subtracts rather than adds value.[8] The net loss is about one-tenth of a percentage per month, that is, roughly one percent per year, so it is quite likely that before trading expenses and management fees the performance was about zero or even mildly positive.

However—and comforting to the fund manager willing to ignore the ample prior evidence on negative alphas—in this study there are doubts about the statistical significance of the negative net outcome: the t-test on the average is only marginally pessimistic ($t(\alpha) = -1.88$), and while the t of the t's is downright negative ($t(t(\alpha)) = -2.53$), it is likely to overstate the significance of the estimate. The results by managing group are similar. Specifically, when we focus on the groups that run at least 30 funds and, therefore, make it possible to do some reliable statistical work (Table 2), we find that not a single group has a statistical significant positive estimated alpha, and two out of six have a substantially negative $t(t(\alpha))$, again making it unlikely that the typical fund has a positive alpha. But also these t-tests overstate the significance. The regressions on the two alternative fund-of-funds, lastly, show alpha's that are more negative in absolute size ($\alpha = -0.261$ and -0.146 percent per month, Table 1) but nevertheless statistically unclear.

Before we proceed, a brief comment on Table 2. The table shows also that the worst performer in terms of $t(t(\alpha))$, group 2, is easily beaten by group 3 when we look at the actual size of the underperformance. That is, the significance here also has to do with the comparatively high precision of the estimates which, in turn, stems from the high amount of diversification: group 2 has the highest average R^2 across management companies. Still, one would expect egregious alphas especially among loose-gun funds that seem to bear little relation to market-wide factors. When we look at that relation in general, there indeed turns out to be a statistically unambiguous negative relation across funds between the square of alpha and R-squared: both the extremely high- and low-alpha funds tend to have low R-squareds (that is, they take on lots of risks unexplained by market, size, or momentum), while the middle-of-the-road performers tend to take on far less stock- or niche-specific risks. We hasten to add that this negative relation is less strong than we had expected (the correlation between α^2 and R^2 is just -0.11, and so is the correlation between $\mid \alpha \mid$ and R^2.)

Let's return to the alphas in Table 1. The mean alphas of about one percent per year

[8] This is not to say that the shareholder is worse off than if (s)he had invested directly into the market.

Table 2: Summary statistics for the largest managers' market, SB, and momentum sensitivities and their average abnormal returns

$$\tilde{r}_f = \alpha_f + \beta_f \tilde{r}_{M,t} + \gamma_f \tilde{SB}_t + \delta_f \tilde{Mo} + \tilde{\epsilon}_{f,t}.$$

	α(%)	β	γ	δ	t(α)	t(β)	t(γ)	t(δ)	$\bar{R}^2$
	manager 1								
avg	-0.043	1.03	0.41	-0.12	-0.04	4.70	1.19	-0.61	0.50
std	1.044	0.62	0.58	0.26	1.08	2.83	1.40	1.61	0.24
t(av)	-0.276	11.00	4.67	-3.18	-0.26	11.04	5.64	-2.53	—
	manager 2								
avg	-0.244	1.21	0.27	-0.08	-0.33	7.30	0.99	-0.62	0.62
std	0.943	0.70	0.50	0.27	0.93	4.76	1.53	1.69	0.26
t(av)	-1.657	11.09	3.50	-1.97	-2.28	9.82	4.13	-2.36	—
	manager 3								
avg	-0.420	1.21	0.59	-0.12	-0.43	5.50	1.83	-0.83	0.53
std	1.499	0.62	0.51	0.26	1.21	2.47	1.35	1.76	0.19
t(av)	-1.610	11.18	6.66	-2.68	-2.03	12.76	7.77	-2.70	—
	manager 4								
avg	-0.079	1.06	0.59	-0.04	-0.05	4.44	1.52	0.01	0.45
std	1.029	0.75	0.55	0.27	0.82	2.99	1.28	1.57	0.23
t(av)	-0.566	10.42	7.88	-1.08	-0.43	10.90	8.71	0.05	—
	manager 5								
avg	0.169	1.01	0.52	-0.02	0.07	3.29	1.52	-0.31	0.56
std	1.136	0.40	0.32	0.27	1.02	2.07	1.11	1.46	0.28
t(av)	0.866	14.77	9.37	-0.52	0.40	9.26	7.98	-1.22	—
	manager 6								
avg	0.106	1.23	0.51	-0.14	-0.03	6.16	1.72	-1.13	0.57
std	1.082	0.54	0.44	0.13	1.10	3.29	1.47	0.99	0.16
t(av)	0.393	9.09	4.67	-4.59	-0.12	7.49	4.67	-4.55	—

Key to Table 2. Monthly excess returns, 1999-2001, on 322 Belgian-managed internationally oriented stock funds are regressed on the world market excess return r_m (MSCI), a small-minus-large return factor SB (DJ STOXX Small minus DJ STOXX Large), and a momentum return Mo (the 25% best winners *vs.* the 25% worst losers over the past six months, taken from a sample of over 2500 individual European stocks). β, γ, and δ estimate the fund's sensitivity to these factors. α estimates the average return that is not explained by any of these factors. Avg stands for average, std for standard deviation, and $t(av)$ for $(avg/std) \cdot \sqrt{N}$.

Table 3: Frequency distributions of estimated parameters and t ratios, 322 funds

$$\tilde{r}_f = \alpha_f + \beta_f \tilde{r}_{M,t} + \gamma_f \tilde{SB}_t + \delta_f \tilde{Mo} + \tilde{\epsilon}_{f,t}.$$

parameter estimates

class	α(%)	freq	cuml	β	freq	cuml	γ	freq	cuml	δ	freq	cuml
1	-3.477	1	0.6	-0.54	0	0.3	-0.97	1	0.6	-0.91	1	0.6
2	-2.865	3	1.6	-0.26	1	0.6	-0.77	3	1.6	-0.79	1	0.9
3	-2,253	4	2.8	0.02	6	2.5	-0.58	0	1.6	-0.68	5	2.5
4	-1.641	12	6.5	0.31	17	7.8	-0.38	0	1.6	-0.57	4	3.7
5	-0.196	26	14.6	0.59	23	14.9	-0.19	10	4.7	-0.46	4	5.0
6	-0.416	64	34.5	0.88	72	37.3	0.01	33	14.9	-0.34	17	10.3
7	0.196	109	68.3	1.16	91	65.5	0.20	57	32.6	-0.23	27	18.6
8	0.808	59	86.7	1.45	41	78.3	0.40	71	54.7	-0.12	64	38.5
9	1.420	21	93.2	1.73	37	89.8	0.59	40	67.1	0.00	67	59.3
10	2.032	13	97.2	2.01	9	92.6	0.79	37	78.6	0.11	61	78.3
11	2.644	4	98.5	2.30	7	94.7	0.99	27	87.0	0.22	49	93.5
12	3.256	2	99.1	2.58	8	97.2	1.18	22	93.8	0.33	9	96.3
13	3.869	2	99.7	2.87	7	99.4	1.38	8	96.3	0.45	2	96.9
14	4.481	0	99.7	3.15	0	99.4	1.57	2	96.9	0.56	4	98.1
15	5.053	0	99.7	3.43	0	99.4	1.77	3	97.8	0.67	2	99.1

t statistics

class	t(α)	freq	cuml	t(β)	freq	cuml	t(γ)	freq	cuml	t(δ)	freq	cuml
1	-2.80	0	0.3	-0.34	5	1.9	-2.18	0	0.3	-5.13	1	0.6
2	-2.33	1	0.6	1.10	20	8.1	-1.72	3	1.2	-4.60	1	0.9
3	-1.87	10	3.7	2.54	61	27.0	-1.25	2	1.9	-4.06	2	1.6
4	-1.40	20	9.9	3.98	50	42.6	-0.78	9	4.7	-3.53	2	2.2
5	-0.93	30	19.3	5.42	49	57.8	-0.32	12	8.4	-2.99	13	6.2
6	-0.46	48	34.2	6.86	50	73.3	0.15	28	17.1	-2.46	10	9.3
7	0.00	73	56.8	8.30	44	87.0	0.61	39	29.2	-1.92	16	14.3
8	0.47	67	77.6	9.74	17	92.2	1.08	42	42.2	-1.39	30	23.6
9	0.94	36	88.8	11.18	9	95.0	1.55	53	58.7	-0.85	42	36.7
10	1.41	19	94.7	12.62	6	96.9	2.01	35	69.6	-0.32	44	50.3
11	1.87	8	97.2	14.06	2	97.5	2.48	29	78.6	0.22	53	66.8
12	2.34	3	98.1	15.50	3	98.5	2.94	30	87.9	0.75	34	77.3
13	2.81	2	98.8	16.94	1	98.8	3.41	16	92.9	1.29	23	84.5
14	3.28	1	99.1	18.38	3	99.7	3.87	13	96.9	1.82	27	92.9
15	4.74	2	99.7	19.82	0	99.7	4.34	4	98.1	2.89	6	99.7

Key to Table 3. Monthly excess returns, 1999-2001, on 322 Belgian-managed internationally oriented stock funds are regressed on the world market excess return r_m (MSCI), a small-minus-large return factor SB (DJ STOXX Small minus DJ STOXX Large), and a momentum return Mo (the 25% best winners *vs.* the 25% worst losers over the past six months, taken from a sample of over 2500 individual European stocks). β, γ, and δ estimate the fund's sensitivity to these factors. α estimates the average return that is not explained by any of these factors, and is in percent per month. Freq stands for absolute frequency, cuml for cumulative relative (percentage) frequency.

hide a degree of variety that would have astounded an investor in the 70s. From Table 1, the range between the best and worst performer is a whopping ten percent per month—and this is a difference between average monthly returns, not month-by-month or *per annum* ones. From Table 2, about one fund in seven has an *ex post* alpha that differs from zero by at least 1.5 percent per month, that is, about 20 percent *per annum*. In line with this, R^2s are now much lower than they used to be. The average is 50 percent, while R^2s in e.g. Ooms, Sercu and Vanthienen (1976) are all above 70 percent and occasionally reach 90 percent. It is equally a sign of the times that our estimated betas range between minus unity and plus four, a far cry from the 0.65-1.1 range found in *e.g.* Ooms, Sercu and Vanthienen (1976). The factor sensitivities are the topic of the next subsection.

2.3 Empirical results (2): exposures & styles

Despite the wide range of estimated market sensitivities (β), the mean is close to unity (Table 1: 1.080), as one would expect. The fact that the average β actually somewhat exceeds unity is due to right-skewness (the presence of a few extremely large numbers): the median beta in Table 1, 0.996, implies that very close to half of the betas are below unity; and about ten percent of the funds have betas exceeding 1.73 (Table 3). The right-skewness probably means that some funds were heavily overweight in the boom-and-bust ICT sectors.[9] Above-average market sensitivity is a characteristic of the two worst-performing groups (groups 2 and 3 in Table 2), but this may be a fluke: group 6, with the highest average market exposure, actually has an above-average alpha.

The average exposure to small-minus-big, gamma, is estimated to be positive (Table 1: 0.442), and quite significantly so. That is, the average fund goes for smaller firms. We also find this for each of the six individual fund-management groups reported in Table 2. Table 3 shows that about 85 percent of the firms have positive gammas; and for the minority, with negative gamma estimates, the significance is dubious. While there is no doubt about the pervasiveness and significance of positive exposure to the small-firm effect, it is far from clear what this positive gamma means: it could be a standard home bias (Belgian firms are small by international standards), or an overweight in small dotcoms, or a conservative

[9] To some extent the extreme betas could be caused by multicollinearity between R_m and SB (ρ=0.75). But that seems to be no major problem: when we compare the betas in Table 1 to the one-factor betas in Table 4, we find that the latter, even though free from multicollinearity-induced imprecision, actually have a somewhat higher cross-sectional standard deviation.

underweight in the (then) bloated ICT giants, or a predilection for a size class where more inefficiencies can be expected. More research is called for here. At any rate, the finding belies the conventional wisdom that most funds are index-chasers or obsessed with liquidity.

Exposure to momentum, lastly, is negative on average (Table 1: -0.059). The effect is, however, the least pervasive of all the exposure patterns discussed so far. Although the t-ratio for the general average is significant (Table 1: -4.218), in Table 2 we find a clearly negative mean for four out of six management groups only; and in Table 3 a mere fifteen percent of the individual deltas are significantly negative by themselves. The mixed picture may mean that behaviour is different across funds. However, at least part of the explanation is that the momentum return is a much lower-impact signal than the market return or the SB factor; and low-impact signals are, of course, hard to pick up and nail down precisely. The properties of the SB and momentum factors are the subject of the next section.

2.4 Empirical results (3): do SB and Mo affect benchmark returns?

Our finding, in the previous section, that funds are exposed to additional factors beside the market return does not in itself mean that fund ratings obtained from a three-factor CAPM would differ systematically from ratings derived from a one-factor CAPM: average returns on the additional factor portfolios may be zero in the sample period. A more indirect mechanism by which the introduction of factors could affect normal returns is that the multivariate beta may differ from the univariate one, provided, of course, that the market risk premium is non-zero. Mathematically, this happens if and only if the new factors are not market-neutral. True, the SB and Mo variables are zero-investment portfolios, or swaps if you like, with both the long and a short leg fully invested in stocks. However, in the past, small stocks tend to have higher betas than large ones, so SB is likely to be not quite market-neutral after all.[10] Let's explore each aspect.

Table 4 shows some relevant summary statistics on the three factors. Strikingly, none of the three factors has a mean return that is unambiguously different from zero; the market return is slightly negative, momentum-chasing did not pay either, and small-v-big yielded only marginally positive results. Thus, in this (small) sample the two mechanisms that may

[10]There also is an uncertain impact on the power of the t-test on alpha. On the one hand the inclusion of more factors lowers the residual standard error which, in itself, increases the power of the t-test on α. On the other hand, the estimation uncertainty about two more regression slopes reduces the precision of the estimated intercept.

Table 4: Three- versus One-factor performance evaluation

	distribution of factors			$\tilde{r}_f = \alpha_f + \beta_f \tilde{r}_{M,t} + \tilde{\epsilon}_{f,t}$			
	r_m	SB	Mo	α	$t(\alpha)$	β	$t(\beta)$
min	-10.803	-9.729	-21.418	-5.861	-6.61	-0.91	-2.79
med	-1.967	-2.528	-2.892	-0.13	-0.21	0.95	5.43
max	7.104	14.889	14.060	26.247	7.49	8.01	26.31
avg	-0.609	1.079	-2.249	-0.104	-0.23	1.05	6.33
std	4.720	5.763	7.789	1.919	1.140	0.678	4.356
t	-0.76	1.11	-1.71	-0.97	-3.67	27.84	26.18
ρ_1	-0.06	-0.10	0.21				
β	1.00	-0.87	0.08				

Key to table 4: The left-hand-side panel of the table summarizes the distributions of the three factors (lowest, midway, and highest value, average, standard deviation, t=$avg/std * \sqrt(N)$, first-order correlation (ρ_1) and the betas of SB and Mo. The factors are the world market excess return r_m (MSCI), a small-minus-large return factor SB (DJ STOXX Small minus DJ STOXX Large), and a momentum return Mo (the 25% best winners *vs.* the 25% worst losers over the past six months, taken from a sample of over 2500 individual European stocks). The right-hand side provides statistics on estimated alphas and betas in a standard one-factor performance evaluation: monthly excess returns, 1999-2001, on 322 Belgian-managed internationally oriented stock funds are regressed on the world market excess return r_m. All returns in the first panel, and the alphas, are in percent per month.

make a three-factor evaluation different from a one-factor version were to a large extent neutralized. Not surprisingly, then, when we look at the mean and median of the one-factor alpha we see almost no difference relative to the three-factor counterparts in Table 1. The net effect on the t-tests, lastly, is mixed: $t(\alpha)$ improves, but $t(t(\alpha))$ weakens. But we still find a significant underperformance after trading costs and other management expenses.

Should we conclude that, in general, adding SB and Mo or other factors will make no difference? Probably not. First, the beta of SB is an impressive -0.87, implying that if the market return had been clearly non-zero—which it is, on average, in larger samples—the market factor would have been substantially confused with the SB one.[11] For example, in the one-factor evaluations there is one outlier fund with a 26 percent alpha, resoundingly t-rated at 7.49. The three-factor evaluation realizes that this is to a large extent a SB effect, and substantially lowers the alpha estimate. This is, of course, an extreme individual case. However, even on average there will, normally, be effects: in the long run the evidence is that small firms do clearly better than large ones, implying that the observed predilection for small firms ought to be accounted for in the benchmark return; and if SB remains not

[11]The negative sign of SB's β is a surprise, though. It explains why our univariate beta is below our multivariate one. On the basis of past international research we would have expected the opposite patterns.

market-neutral, then in a sample with a non-zero market return we should disentangle true market sensitivity from indirect market sensitivity via SB.

3 Conclusion

Looking at a large number of Belgian-managed equity funds, we find that the average estimated alpha is negative to the tune of about (-)1 percent per year, but statistically ambiguously so. Looking at the larger fund-management groups as separate subsamples, we find that most of them similarly score below zero. Our conclusion is independent of whether we apply one- or three-factor evaluations. That last finding, however, is likely to be sample-specific; in larger samples, the conclusions can be expected to differ. The main reason is that SB is far from market-neutral, and funds are clearly loading positively on SB. Thus, when each factor has a positive risk premium, we ought to disentangle the sensitivities. With respect to momentum, the exposures are less clear, but the same argument might apply. And even in our sample, where the average alphas happen to be similar, knowledge of an individual fund's one-factor alpha explains a mere 16 percent of the variance of its three-factor score. In short, the additional factors matter.

Still, our study is exploratory in several respects. First, we have no large-scale survey on what factors were relevant and successful in explaining passive, static style portfolios. Book-to-market, for instance, and exchange rates are missing. Second, we do not know what the funds' positive exposure on SB stands for: home bias (Belgian firms are small), or dot-com gambles, or ICT aversion. Third, it is not obvious what the negative momentum exposure means (provided it holds up in larger samples at all): timely cutting of losses, or premature profit-taking. Lastly, the sample period (confined to the still brief life of DJ STOXX) is quite short.

References and further reading

Carhart, M. (1997), On Persistence of Mutual Fund Performance, *Journal of Finance* 52, 57-82

Crombez, J. (2001), *Modeling Expected Returns and Discount Factors for European Stock Markets*, PhD dissertation, RU Gent

Fama, E.F., and K. French (1992), The Cross-Section of Expected Stock Returns, *Journal*

Ooms, M., Sercu, P. and Vanthienen, L. (1976), The Performance of Belgian Mutual Funds, *Tijdschrift voor Economie en Management* 21, 1967

Merton, R. C. (1973), An Intertemporal Asset Pricing Model, Econometrica 41, 867-887

Sercu, P. (1980): A Generalization of the International Asset Pricing Model, *[Journal de l'Association Francaise de] Finance* **1**(1), 91-135

Sercu, P. (2000), Methodological Issues and Advances in Mutual-Fund Performance Evaluation, *Revue de la Banque/Tijdschrift voor Bank- en Financiewezen* 9, 558-567

Sharpe, W.F. (1964), Capital asset prices: a theory of market equilibrium under conditions of risk, *Journal of Finance* 10, 425-442

Sharpe, W.F. (1992), Asset Allocation: Management Style and Performance Measurement, *Journal of Portfolio Management*, 7-19

Wermers, R. (1999), Mutual Fund Herding and the Impact on Stock Prices, *Journal of Finance* 54, 581-622